UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPRDC TR 86-17	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Navy Personnel Research and Development Center 6b OFFICE SYMBOL (If applicable) Code 62	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) San Diego, CA 92152-6800	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING Office of the Assistant Secretary of Defense 8b. OFFICE SYMBOL (If applicable) MI&L	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Washington, DC 20301-4000	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
MIPR-USAF			83-4

11 TITLE (Include Security Classification)
VALIDITY GENERALIZATION OF NAVY SELECTOR COMPOSITES

12. PERSONAL AUTHOR(S)
Foley, Paul P.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Report	FROM 82 Oct TO 84 Oct	1986 May	40

16 SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	09	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Test validation, selection and validation, validity generalization, ASVAB

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Validation of the Armed Services Vocational Aptitude Battery (ASVAB) has been considered necessary for each Class "A" school and with introduction of each new ASVAB form. Scientific litrature has presented evidence that variability in observed validity coefficients across studies may be attributable to a numbr of fctors and the need for revalidation of the ASVAB may be unnecessary. If validity coefficients could be generalized from one Navy Class "A" school to a number of related schools, the Navy could save substantial costs in revalidating the ASVAB.

Results obtained from three sets of analyses confirmed the generalizability of validity coefficients across a wide range of Class "A" schools for each of the four selector composites investigated. The results will be used to estimate validity coefficients of the ASVAB in the future thus eliminating the cost of a new study. Validity generalization results obtained were comparable to those reported in the literature for similar studies.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Foley, Paul P.	22b. TELEPHONE (Include Area Code) (619) 225-2176 22c. OFFICE SYMBOL Code 62

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

NPRDC-TR 86-17

MAY 1986

5-1-91

VALIDITY GENERALIZATION OF NAVY SELECTOR COMPOSITES

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA

NAVY PERSONNEL RESEARCH
AND
DEVELOPMENT CENTER
San Diego, California 92152

UNITED STATES NAVY
NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER

NPRDC TR 86-17 May 1986

VALIDITY GENERALIZATION OF NAVY SELECTOR COMPOSITES

Paul P. Foley
Leonard Swanson
Stephanie Booth-Kewley

Reviewed by
John J. Pass

Approved by
Martin F. Wiskoff

Released by
Howard S. Eldredge
Captain, U.S. Navy
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152-6800

FOREWORD

This research was conducted within MIPR AFMPC-83-4 (Armed Services Vocational Aptitude Battery (ASVAB-DoD)), which was funded by the Office of the Assistant Secretary of Defense (MI&L). The overall purpose of the research was to investigate the application of new technology to ASVAB for establishing selection and classification standards, developing new test forms, and validating them against school and on-job performance measures. The specific objectives of the present research were to (1) investigate the utility of validity generalization for estimating future validity coefficients within specific Navy ratings, and (2) assess the extent to which similarities between jobs within the same family reduce the variability of validity coefficients. This study of validity generalization procedures for Navy jobs can be applied to data from the other services to identify common selector composites. Results are intended for use by organizations conducting ASVAB research and by Naval Military Personnel Command and Navy technical school personnel.

This report is the third in a series conducted under this work unit. The first (NPRDC SR 83-4) described the development of a deliberate failure key for the Armed Forces Qualification Test (AFQT) portion of ASVAB. The key was designed to identify persons who, in the event of resumption of the draft, would attempt to fail the AFQT in order to avoid military service. The second report (NPRDC TR 85-19) compared univariate and multivariate corrections for range restriction to determine which would be the most accurate for use in ASVAB validation.

Appreciation is expressed to Dr. Kenneth Pearlman, formerly of the Office of Personnel Management and now with American Telephone and Telegraph, for providing the computer program used in the data analysis, and to Mr. Greg Candell for modifying the program and for conducting the various analyses.

H. S. ELDREDGE
Captain, U.S. Navy
Commanding Officer

J. W. TWEEDDALE
Technical Director

SUMMARY

PROBLEM

Personnel psychologists have traditionally considered it necessary to validate the Armed Services Vocational Aptitude Battery (ASVAB) for each Class "A" school and for each new test form. In addition, they perform validation studies as needed by particular schools for specific reasons, such as a sudden increase in student attrition or a curriculum change. However, the recent scientific literature has presented strong evidence that most of the variability in validity coefficients among various studies is artifactual. That is, the variability in observed validity coefficients across studies may be due to factors such as sampling error, criterion unreliability, test unreliability, and restriction in range of ability. The long-standing belief in the situational specificity of validity coefficients has been called into question. If predictive validity coefficients, rather than being specific, could be generalized across a wide range of jobs, tasks, and situations, the Navy could save substantial costs in revalidating the ASVAB.

OBJECTIVES

The objectives of this research were to determine whether validity coefficients are generalizable (1) across studies conducted within individual Navy ratings, (2) across all Navy ratings, and (3) across ratings within systematically formed rating families. A secondary objective was to assess the usefulness of different systems for forming families of ratings.

APPROACH

The present study used the Basic Test Battery (BTB), predecessor of the ASVAB, because more validity data were available for the BTB. The research analyzed validity coefficients of four BTB selector composites used for predicting final school grades in 90 Class "A" schools. To investigate the generalizability of these coefficients, three analyses were conducted: (1) on distributions of six or more validity coefficients that had been calculated for the operational BTB composite in 16 schools, (2) on validity coefficients for all Navy ratings, and (3) on validity coefficients obtained for ratings that had been systematically formed into families. Two validity generalization procedures were compared, one in which the mean of each distribution of validity coefficients was corrected for range restriction and criterion unreliability, and one in which the validity coefficients were individually corrected.

The utility of a rating family for classification purposes was inferred by the extent to which it reduced variability of validity within systematically formed rating families as compared to all Navy ratings combined and to randomly formed rating families. It was hypothesized that variability would be smaller for systematically formed rating families.

RESULTS AND DISCUSSION

Results obtained from the three sets of analyses confirmed the generalizability of validity across a wide range of jobs for each of the four BTB composites. Validity generalization results obtained using two different procedures were comparable to those reported in the literature for BTB tests. Correcting validity distributions for sampling error alone (mean observed validities corrected for range restriction) is useful because validity coefficients individually corrected for restriction in range are not generally available.

The average variability of validity coefficients for systematically formed rating families was smaller than for all Navy ratings and for randomly formed rating families, as hypothesized. The differences, however, were not great in the sense of significantly moderating the variability of validity coefficients, nor was any one grouping system uniformly more effective than any other.

CONCLUSIONS AND RECOMMENDATIONS

The limited effect of the three systematically formed rating families in moderating the variability implies that a simplistic strategy for grouping ratings into families may be as effective, in terms of validity generalization results, as the more expensive and complicated task analysis.

The apparent similarity of the BTB to the ASVAB implies that these procedures will also prove useful in the analysis of ASVAB data. It is recommended that a data base similar to that of the BTB be developed for the ASVAB. Results of applying validity generalization procedures to ASVAB data base would be used to (1) estimate the validity of selector composites before validation data are available, (2) assess the need for a revalidation of selector composites, and (3) suggest ways of combining criterion data from small Class "A" schools to provide samples of sufficient size for validation.

CONTENTS

LIST OF TABLES

INTRODUCTION

Background

All the United States military services use the Armed Services Vocational Aptitude Battery (ASVAB), composed of 10 tests, as their primary instrument to select and classify enlisted military personnel. The services use the Armed Forces Qualification Test (AFQT), a composite of four ASVAB tests, to determine eligibility for enlistment. In addition, the Navy uses 11 different composites (of two, three, or four tests) to assign personnel to Class "A" technical schools and to predict their probability of successfully completing technical training.

Since new forms of the ASVAB were introduced in 1984, it has been necessary for the Navy to validate all its selector composites derived from the new forms. These composites are validated against school performance measures (final school grade in lockstep courses and time-in-training in self-paced courses) for more than 100 Navy "A" schools. Although it is important that the validity of the new ASVAB forms be determined shortly after they are implemented, the process usually takes at least 2 years. Sufficient enlistees must be tested with the new forms and complete recruit and Class "A" school training, and a validation study must be completed. Some Navy schools are never included in routine validation studies because samples from them are too small to permit meaningful analysis. In addition to studies conducted when new ASVAB forms are introduced, ASVAB validation for individual "A" schools is occasionally needed for specific reasons, such as dramatic increases in student attrition rate, changes in school curriculum resulting from changes in the Navy occupation or method of instruction, merging of ratings, and creation of new ratings.

Problem

Validity studies, whether conducted for a large number of Class "A" schools or an individual school, produce costs for the schools providing the performance data, for the Chief of Naval Education and Training who supplies computer files containing ASVAB scores, and for the Navy Personnel Research and Development Center (NAVPERSRANDCEN), where the separate data bases are merged, analyses conducted, and reports documenting the findings produced. Although the average cost of a Navy validation study has not been determined, based on survey data, Outerbridge (1979) estimated the cost of a typical validation study at more than $25,000 and sometimes at more than several hundred thousand dollars.

Personnel psychologists believe that validation and revalidation of the ASVAB for individual Navy schools is required because test validity is situationally specific, that is, a selector composite valid for one "A" school is not likely to be valid for another. Accordingly, empirical validation of selector composites for each individual Class "A" school has traditionally been considered necessary to establish that a selector composite is effective for assignment to that school and, if there has been a change in curriculum, that the selector composite is still valid for the redesigned course. Researchers have observed the substantial variability in validity coefficients of predictor tests across studies, even across studies for the same school or job and selection test.

Schmidt and Hunter (1977) and their colleagues have presented strong evidence that most of this variability is due to statistical artifacts. Accordingly, they have developed validity generalization analysis as a method for determining the extent to which "true" validity coefficients, those with their artifactual variance removed, vary across studies.

They identified the four major sources of artifactual variance as (1) sampling error, (2) differences among studies in criterion unreliability, (3) differences in test reliability, and (4) differences in degree of range restriction. They identified three additional sources of artifactual variance, (1) differences in criterion contamination and deficiency, (2) computational and typographical errors, and (3) differences in factor structure among tests purporting to measure the same construct. Validity generalization analysis consists of correcting the observed variability in validity coefficients for the first four sources of artifactual variance; there is no adequate way to correct for the latter three.

In the terminology of Schmidt and Hunter, combining the first four sources of artifactual variance yields the predicted variance, which, when subtracted from the total observed variance, in turn yields the residual variance. If residual variance is essentially zero, the hypothesis of situational specificity may be rejected and validity generalization accepted. The mean and standard deviation (SD) of the distribution of true validity coefficients are obtained by correcting the mean and SD of this residual distribution for criterion unreliability and restriction in range. For each distribution, the estimated SD of the distribution of true validities is multiplied by 1.2816, that point on the abscissa of the normal curve below which 90 percent of the area lies, and this value is subtracted from the estimated mean true validity. One can then conclude, with 90 percent confidence, that true validity is at or above this value.

Schmidt, Hunter, and their colleagues conducted many studies (e.g., Pearlman, Schmidt, & Hunter, 1980; Schmidt, Hunter, & Caplan 1981) indicating that 70-80 percent of the observed variance in validity coefficients can be accounted for by the four statistical artifacts for which it is possible to correct. Their findings led them to conclude that differences among jobs, tests, and organizational settings moderate validity coefficients much less than had previously been thought.

Pearlman (1982) also studied the effects of different job grouping strategies on validity generalization outcomes data taken from the predecessor to ASVAB, the Navy Basic Test Battery (BTB). He used validity coefficients for six BTB tests (general classification, arithmetic, clerical, mechanical comprehension, shop practices, and electronics technician selection), computed for 500 school samples in 61 Navy ratings. Four types of job descriptors were used to group ratings into families. Pearlman found in every case that a test valid for some ratings in a family was also valid for the others. These results confirmed earlier findings that suggested that test validities are generalizable not only for individual Navy "A" schools and ratings, but also across a large sampling of "A" schools representing a wide variety of ratings.

Validity generalization has several implications for the Navy's validation work on the ASVAB. It could provide the rationale for combining small Class "A" schools with others in related occupational fields to obtain samples large enough for meaningful analysis. In addition, validity generalization procedures could provide evidence of validity for a new test known to be parallel or similar to an operational test, and thereby justify implementation of the new test before validation. Finally, if validities generalize across ratings, there would be less need for separate validation of each "A" school; research resources could be reallocated to other problems of the Navy personnel and classification system.

Objectives

The objectives of this research were to determine whether validity coefficients are generalizable (1) across studies conducted within individual Navy ratings, (2) across all

Navy ratings, and (3) across ratings within systematically formed rating families. A secondary objective was to assess the usefulness of different systems for forming families of ratings.

APPROACH

Data Sources

This research used the data analysis procedures developed by Pearlman and his colleagues (Pearlman et al., 1980; Pearlman, 1982). Pearlman (1982) investigated validity generalization for single BTB tests; the present research used validation data for four BTB selector composites used to make school assignments. Data were taken from the same sources that Pearlman used: Naval Personnel Research Activity (1961), Thomas and Thomas (1965, 1967), Thomas (1970), and Swanson (1977).

Variables

Predictors

The predictor variables consisted of the four BTB selector composites, general technical (GT), mechanical (MECH), electronics (ELEC), and clerical (CLER), each of which is composed of two or three tests (see Table 1). Because of the high similarity of the general classification and clerical tests across forms, validity information for GT (composed of the general classification and arithmetic tests) and CLER (composed of the general classification and clerical tests) was obtained from both Forms 6 and 7 of the BTB. The arithmetic tests, although not identical across forms, were considered sufficiently similar to be treated as the same test. Validity information for MECH and ELEC was drawn exclusively from BTB Form 7, because the shop practices and electronics technician selection tests were not included in Form 6. A description of the test content is provided in Appendix A.

Criterion

The criterion for each validation study was final school grade (FSG). It varied from school to school, but was generally based on periodic quizzes, performance measures, and a final examination. Grades ranged from 63 to 99.

Samples

Validity coefficients from 90 ratings were compared for GT, MECH, ELEC, and CLER (see Appendix B). Three different subsets of those 90 ratings were used to investigate validity generalization:

1. Across studies within individual Navy ratings.

2. Across all Navy ratings.

3. Within systematically formed rating families.

The first subset was comprised of ratings that had six or more validity coefficients for the operational composite used for selection. Ratings that had validity information from BTB Form 6 for the MECH and ELEC composites were not used. The second subset

Table 1

Selection Composites in Basic Test Battery (BTB), Forms 6 and 7

Composite	Tests Included in Composite Form 6	Tests Included in Composite Form 7
General Technical (GT)	GCT + ARI	GCT + ARI
Mechanical (MECH)	GCT + MECH	GCT + MECH + SP
Electronics (ELEC)	GCT + ARI + ETST	ARI + 2ETST
Clerical (CLER)	GCT + CLER	CGT + CLER

Notes:

1. Test abbreviations and names are:

   GCT = general classification test
   ARI = arithmetic
   MECH = mechanical
   CLER = clerical aptitude
   SP = shop practices
   ETST = electronics technician selection test

2. ETST was not included in BTB Form 6; at that time it was a special test administered only to applicants for electronics schools.

included all Navy ratings except for those that could not be placed into occupational groups. The third subset was composed of ratings which could be placed into one of the occupational groups of the three systematically formed rating families. Some ratings could not be placed into occupational groups and were eliminated from the systematically formed rating families as well as the all Navy rating subset.

Analyses of Validity Generalization

Validity generalization was investigated by two procedures that estimate the mean and SD of distributions of true validities: mean corrected validities (MCV) and individually corrected validities (ICV). In the first procedure, the mean of each distribution was corrected for range restriction and criterion unreliability. In the second, the validities were individually corrected for these sources of error. The MCV procedure corrects observed variance for sampling error only, while the ICV procedure corrects for range restriction differences among studies, as well. The purpose of using both was to evaluate the MCV procedure for use in situations where data necessary for the ICV procedure are not available. An explanation of each procedure and worked examples are provided in Appendix C.

Within Ratings

For each of the 16 ratings used, a distribution of coefficients was drawn from at least six validity studies (see Appendix D). Validity generalization was investigated for ratings that used both the MCV and ICV procedures. Composite validities were considered generalizable across studies if the 90-percent-credibility value was substantial and exceeded zero.

Across All Navy Ratings and Within Rating Families

To investigate validity generalization across all Navy ratings, validities for the ratings were pooled into one heterogeneous group. To investigate validity generalization across ratings within families, ratings were grouped using three different job family systems (see Appendix E). These job family systems were based on (1) the 9 occupational groups of the Navy Occupational Handbook (NOH, U.S. Navy, 1966), (2) the 20 occupational groups of the Navy's official classification manual (CLASSMAN, U.S. Navy, 1975), and (3) the 4 occupational groups formed according to the BTB selector composites (BTBFAM) used for "A" school assignment. For a rating family system to be useful for selection, the variability in validity coefficients within a systematically formed family should be smaller than that observed across all Navy ratings and smaller than the variability observed within randomly formed groups of ratings. Accordingly, and based on previous related research (Pearlman, 1982), it was hypothesized that the estimated true SDs of validities observed within rating families formed using job family systems would be smaller than the SDs for all ratings combined and for those within comparable randomly formed rating families. However, these differences would not significantly moderate validity variability, that is, produce different validity generalization conclusions for rating families than for all ratings combined.

A series of random rating families were formed so that each randomly formed job family system had the same number of groups as each of the systematically formed rating family systems. For example, the NOHFAM rating family system is made up of nine occupational groups, so the corresponding random system was also made up of nine groups. The number of groups in an analysis of variance influences the within-group variance; the within-group variance decreases in relation to the between-group variance as the number of groups is increased. Making the number of groups the same for the systematically and randomly formed job family systems made it possible to observe the effects of different substantive systems on validity.

RESULTS AND DISCUSSION

Validity Generalization Within Ratings

As shown in Table 2, results yielded by the two different validity generalization analysis procedures were highly similar across studies conducted within ratings. This result was especially evident from comparison of the sample weighted means for the two procedures. Substantial mean true validities and 90-percent-credibility values were obtained using both validity generalization procedures for all 16 of the ratings. In no instance did the credibility values include zero. The mean true validities ranged from .33 to .87 for the MCV procedure, and from .33 to .78 for the ICV procedure. The 90-percent-credibility values ranged from .24 to .74 for the MCV procedure, and from .21 to .72 for the ICV procedure. One can be 90 percent certain that in future validation studies, the validity coefficient for the SM rating, for example, will not be zero nor will it be lower than .24 (MCV) or .21 (ICV). The best estimate of the validity in a future study is .33.

Table 2

Statistical Summary for Distributions of True Validity Coefficients of BTB Composites in 16 Ratings

Rating[a]	N Samples	N Students	Observed Selector Composite Validity Mean	Observed Selector Composite Validity SD	Validity Generalization Based on Mean Corrected Validity (MCV) Mean	MCV SD	MCV 90% CV[b]	Validity Generalization Based on Individually Corrected Validity (ICV) Mean	ICV SD	ICV 90% CV[b]
QM	6	2,783	.41	.073	.50	.075	.41	.55	.003	.55
SM	8	2,885	.24	.072	.33	.070	.24	.33	.096	.21
FT	7	2,952	.56	.072	.87	.101	.74	.78	.048	.72
TM	6	1,357	.29	.123	.47	.176	.25	.46	.172	.24
YN	12	5,036	.34	.090	.47	.109	.33	.49	.085	.38
SK	10	8,153	.44	.061	.58	.071	.49	.57	.043	.51
DK	10	1,624	.32	.078	.43	.047	.37	.43	.093	.31
EM	6	8,469	.42	.049	.61	.062	.53	.61	.065	.53
IC	6	4,038	.44	.076	.63	.098	.51	.60	.059	.53
HM	12	34,471	.58	.039	.73	.046	.67	.72	.035	.67
ET	9	9,698	.49	.053	.82	.083	.72	.72	.042	.67
ETR	7	4,865	.42	.119	.74	.202	.48	.60	.103	.47
ETN	8	6,079	.43	.091	.75	.150	.55	.62	.079	.52
OS	12	9,099	.41	.091	.65	.136	.47	.64	.064	.56
TM	8	1,463	.38	.049	.56	.000	.56	.55	.086	.44
PH	6	2,418	.49	.050	.64	.042	.59	.64	.000	.64
Total	133	105,390								
Weighted means[c]			.42	.073	.61	.090	.49	.58	.067	.50

[a]See Appendix B for full names of ratings.

[b]Credibility value.

[c]Means weighted according to sample size.

As shown in Table 3, the estimated true validity coefficients for each of the four composites ranged from .43 for the CLER composite to .69 for the ELEC composite based on MCV, and from .53 to .62 respectively based on ICV. The 90-percent-credibility values ranged from .28 to .51 for the MCV procedure and from .40 to .48 for the ICV procedure. These results support validity generalization for BTB composites across the diverse ratings used in the analyses. In all instances the "true validity" was in a positive range; none of the 90-percent-credibility values included zero. These results support generalizability to other Navy jobs from which these ratings were drawn.

The findings in Table 3 also support findings from earlier work. The mean true validity, SD, and 90-percent-credibility value (for ICV) averaged across the four BTB composites (of two or three tests) were .57, .11, and .43, compared to .46, .10, and .33 averaged across six BTB tests reported by Pearlman and Schmidt (1981). Both results were similar to the values of .45, .11, and .30 reported by Schmidt, Hunter and Pearlman (1981), which were obtained by averaging across all the subtests of the Army Classification Battery. Schmidt et al. (1981) corrected the mean true validities for range restriction but not for criterion unreliability.

Table 3

Statistical Summary for Distributions of True Validity Coefficients of BTB Composites Across Navy Ratings

Composite[a]	N			Observed Validity		Validity Generalization Based on MCV			Validity Generalization Based on ICV		
	Ratings	Samples	Students	Mean	SD	Mean	SD	90% CV[b]	Mean	SD	90% CV[b]
GT	82	396	276,847	.43	.106	.61	.142	.43	.60	.098	.48
MECH	77	254	212,972	.38	.101	.55	.139	.37	.55	.118	.40
ELEC	77	203	176,068	.47	.100	.69	.141	.51	.62	.125	.46
CLER	81	391	271,097	.34	.101	.43	.121	.28	.53	.093	.41

[a]See Table 1 for full names of composites.

[b]Credibility value.

Validity Generalization Within Systematically Formed Rating Families

The lowest credibility values for the 120 validity distributions within rating families (see Tables 4, 5, and 6) are .13 and .14 (Table 5, for the MECH composite). Aside from this exception, the credibility values are substantial; in no distribution does the credibility value include zero. These results indicate the generalizability of validity for the four BTB composites across ratings within job families formed using three different job family systems. In addition, validity generalization results based on MCVs were very similar to those based on ICVs.

Usefulness of Different Rating Family Systems

The estimated SDs of true validities were obtained by using the ICV procedure for each of the four BTB composites averaged across the systematically formed rating families (see Table 7). For example, the value for the NOHFAM for the GT composite is .091; this value is the average SD (weighted by N samples) for the nine NOHFAM rating families. The values in parentheses were calculated in the same way for the randomly formed rating family systems. The last row of Table 7 contains the true SDs (see the last row of Table 3) obtained for all Navy ratings combined, while the last column contains the average estimated true SDs averaged across BTB composites. The utility of a rating family for classification purposes is indicated by how much it reduces (1) variability of validity within rating families compared to that observed across all Navy ratings, and (2) variability compared to that observed within randomly formed rating families.

As shown in Table 7, the weighted average SDs of true validities for each of the four BTB composites were smaller for the substantive rating families (NOHFAM, CLASSFAM, and BTBFAM) than for all Navy ratings combined. Overall, however, the differences were rather small. The largest difference was for the MECH composite in the CLASSMAN families (.079 - .118 = -.039), while the smallest difference was observed for the GT composite in the BTBFAM families (.096 - .098 = -.002). The average SDs for the substantive rating families were also smaller than for the corresponding randomly formed families, but overall the differences were small. The largest difference between substantive and random values was for the MECH composite in the BTBFAM system (.097 - .118 = -.021), while the smallest was for the CL composite in the BTBFAM system (.088 - .088 = .000). Thus, these results supported the hypothesis that the SDs averaged across occupational groups of substantive rating families would be smaller than for all Navy ratings combined and the average for groups of random ratings. The differences, however, were not great in the sense of significantly moderating the variability of validity coefficients. Nor was any one grouping system uniformly more effective than any other.

Table 8 shows the data in Table 7 in the form of relative rather than absolute differences. Column 1 for the NOHFAM GT composite shows the variance (the square of corresponding SDs from Table 7) as a percentage of variance for all Navy ratings combined $(.091)^2 / (.098)^2 = .86 \times 100 = 86\%$. Column 2 shows the comparable values for random groups of ratings (.092) / (.098) = .88 X 100= 88%. The difference column represents the reduction in variance attributable to substantive groupings of ratings beyond that which is due to chance.

The last three columns (Means) of Table 8 contain values averaged across BTB composites. For example, in Table 7 the difference between the average SDs for the NOHFAM and all Navy ratings is .016 (.109 - .093), equivalent to a 26 percent (100% - 74%) reduction in variance attributable to the NOHFAM rating system. The corresponding reduction in variance for random groups was 16 percent (100% - 84%). Of the 26

Table 4

Statistical Summary for BTB Composites for Rating Families Based on the Navy Occupational Handbook (NOH)

Composite and Rating Family	N		Observed Validity		Validity Generalization Based on MCV			Validity Generalization Based on ICV		
	Samples	Students	Mean	SD	Mean	SD	90% CV[b]	Mean	SD	90% CV[b]
GT										
1	22	8,146	.45	.152	.58	.190	.34	.54	.149	.34
2	52	14,591	.40	.090	.61	.116	.46	.62	.091	.50
3	42	38,376	.35	.074	.58	.114	.43	.60	.088	.49
4	65	36,159	.40	.086	.56	.108	.42	.53	.091	.41
5	43	43,145	.41	.085	.57	.113	.43	.60	.070	.51
6	39	10,313	.49	.104	.58	.110	.44	.61	.091	.49
7	107	77,883	.42	.086	.61	.117	.46	.60	.092	.49
9	22	47,861	.55	.063	.70	.076	.60	.68	.071	.59
Total	392	276,474								
Weighted means[c]			.42	.089	.59	.116	.45	.59	.091	.48
MECH										
1	17	6,686	.30	.129	.42	.171	.20	.40	.136	.23
2	40	10,326	.36	.078	.56	.086	.44	.57	.087	.46
3	30	32,673	.41	.074	.60	.101	.46	.59	.069	.51
4	43	30,753	.23	.071	.33	.091	.22	.38	.104	.25
5	23	26,542	.39	.073	.58	.102	.45	.60	.058	.52
6	28	9,097	.49	.097	.65	.115	.50	.63	.084	.52
7	61	57,878	.40	.082	.60	.116	.45	.58	.103	.45
9	11	38,785	.42	.036	.56	.044	.50	.57	.057	.50
Total	253	212,740								
Weighted means[c]			.37	.081	.54	.104	.40	.54	.091	.43
ELEC										
1	13	5,318	.45	.117	.64	.155	.44	.53	.101	.40
2	36	9,325	.47	.125	.74	.182	.51	.66	.149	.46
3	30	32,653	.50	.100	.82	.158	.62	.71	.103	.58
4	30	24,993	.43	.077	.61	.102	.48	.53	.094	.41
5	15	13,813	.49	.070	.70	.093	.58	.65	.086	.54
6	24	8,291	.47	.115	.61	.139	.43	.57	.113	.43
7	47	50,939	.47	.090	.69	.127	.52	.61	.111	.47
9	7	30,504	.49	.114	.64	.148	.45	.60	.130	.43
Total	202	175,836								
Weighted means[c]			.47	.100	.69	.139	.51	.61	.112	.47
CLER										
1	22	8,146	.34	.119	.44	.139	.26	.47	.128	.30
2	53	14,719	.30	.087	.40	.091	.29	.53	.069	.44
3	42	38,376	.24	.064	.34	.078	.24	.51	.082	.40
4	62	34,409	.33	.076	.42	.084	.31	.47	.084	.36
5	42	41,982	.31	.065	.40	.076	.31	.52	.057	.45
6	39	10,313	.40	.092	.45	.085	.34	.53	.067	.45
7	106	77,639	.32	.075	.42	.089	.31	.52	.084	.41
9	21	45,140	.48	.038	.55	.039	.51	.63	.067	.54
Total	387	70,724								
Weighted means[c]			.33	.077	.42	.085	.31	.52	.078	.41

[a]Please refer to Table 1 for full names of composites.

[b]Credibility value.

[c]Means weighted according to sample size.

Table 5

Statistical Summary for BTB Composites for Rating Families Based on the Navy Classification Manual (CLASSMAN)

Composite and Occupational Group	N		Observed Validity		Validity Generalization Based on MCV			Validity Generalization Based on ICV		
	Samples	Students	Mean	SD	Mean	SD	90% CV[b]	Mean	SD	90% CV[b]
GT										
1	8	2,885	.30	.107	.40	.126	.24	.37	.112	.23
2	8	3,465	.59	.067	.70	.070	.61	.66	.052	.59
3	38	41,113	.41	.085	.58	.113	.43	.61	.066	.52
4	13	2,538	.34	.109	.47	.126	.30	.48	.108	.34
5	67	62,765	.42	.081	.62	.111	.47	.61	.083	.50
6	13	3,777	.32	.051	.45	.000	.45	.43	.057	.36
7	5	3,000	.40	.058	.60	.071	.51	.57	.075	.48
8	51	44,637	.35	.072	.59	.110	.45	.61	.079	.51
9	33	7,608	.40	.094	.59	.111	.45	.59	.117	.44
10	5	1,340	.37	.054	.61	.000	.61	.54	.048	.47
11	5	1,242	.55	.094	.76	.114	.61	.73	.063	.65
12	3	301	.46	.059	.70	.000	.70	.64	.000	.64
13	39	10,313	.49	.104	.58	.110	.44	.61	.091	.49
14	17	38,909	.49	.058	.71	.070	.62	.70	.051	.64
15	24	8,663	.42	.090	.55	.104	.42	.54	.087	.43
16	29	13,730	.40	.043	.54	.099	.42	.52	.090	.41
17	10	2,796	.48	.071	.64	.071	.55	.63	.030	.59
20	15	9,437	.40	.112	.60	.161	.39	.56	.110	.41
Total	383	258,519								
Weighted means[c]			.41	.081	.59	.102	.45	.58	.082	.48
MECH										
1	6	2,494	.16	.076	.24	.089	.13	.27	.103	.14
2	6	2,783	.42	.041	.54	.020	.52	.52	.058	.45
3	20	25,456	.39	.073	.59	.103	.45	.60	.059	.53
4	7	1,287	.37	.105	.56	.130	.39	.54	.068	.45
5	36	45,889	.42	.071	.63	.101	.50	.61	.090	.49
6	9	3,237	.30	.082	.49	.109	.35	.43	.100	.31
7	3	2,670	.31	.036	.44	.026	.41	.48	.089	.36
8	37	36,162	.40	.073	.59	.099	.46	.60	.065	.51
9	27	6,420	.35	.074	.55	.075	.45	.54	.101	.41
10	4	953	.26	.071	.41	.052	.34	.37	.000	.37
11	3	837	.48	.094	.67	.146	.49	.66	.086	.55
12	3	301	.34	.066	.52	.000	.52	.53	.000	.53
13	28	9,907	.49	.097	.65	.115	.50	.63	.084	.52
14	9	32,750	.42	.034	.57	.041	.51	.59	.046	.53
15	17	8,420	.26	.068	.38	.078	.28	.42	.082	.32
16	18	10,889	.23	.072	.35	.091	.23	.38	.101	.25
17	3	1,696	.47	.010	.60	.000	.60	.60	.000	.60
20	9	7,113	.22	.074	.32	.096	.20	.41	.109	.27
Total	245	199,254								
Weighted means[c]			.37	.073	.54	.090	.42	.54	.079	.44

Table 5 (Continued)

Composite and Occupational Group	N Samples	N Students	Observed Validity Mean	Observed Validity SD	Validity Generalization Based on MCV Mean	Validity Generalization Based on MCV SD	Validity Generalization Based on MCV 90% CV[b]	Validity Generalization Based on ICV Mean	Validity Generalization Based on ICV SD	Validity Generalization Based on ICV 90% CV[b]
ELEC										
1	4	1,968	.34	.052	.46	.048	.40	.40	.000	.40
2	4	1,941	.57	.063	.73	.070	.64	.64	.000	.64
3	13	13,201	.49	.067	.71	.090	.60	.66	.080	.56
4	6	793	.41	.110	.55	.118	.40	.52	.102	.39
5	28	40,671	.47	.083	.70	.119	.54	.62	.103	.49
6	8	3,017	.33	.052	.48	.034	.44	.43	.072	.34
7	2	2,233	.42	.050	.64	.067	.56	.59	.093	.47
8	37	36,162	.51	.099	.82	.156	.62	.72	.101	.59
9	23	5,419	.43	.127	.66	.177	.43	.58	.152	.39
10	4	953	.42	.056	.69	.029	.65	.52	.000	.52
11	2	710	.65	.020	.89	.000	.89	81	.000	.81
12	3	301	.48	.083	.72	.055	.65	.65	.000	.65
13	24	8,291	.47	.115	.61	.139	.43	.57	.113	.43
14	6	27,229	.50	.120	.65	.156	.45	.61	.135	.44
15	10	4,966	.42	.057	.59	.060	.51	.53	.049	.47
16	13	9,048	.42	.054	.59	.062	.51	.51	.068	.43
17	2	1,377	.52	.015	.68	.000	.68	.65	.000	.65
20	7	6,404	.52	.054	.73	.068	.64	.65	.047	.59
Total	196	164,684								
Weighted means[c]			.46	.087	.68	.114	.53	.60	.090	.49
CLER										
1	8	2,885	.24	.072	.33	.070	.24	.33	.096	.21
2	8	3,465	.44	.097	.53	.107	.39	.56	.072	.47
3	37	39,950	.31	.067	.41	.079	.30	.53	.055	.45
4	13	2,538	.28	.095	.36	.089	.25	.41	.086	.30
5	65	62,223	.32	.068	.42	.081	.32	.53	.078	.43
6	13	3,777	.24	.046	.30	.000	.30	.38	.028	.35
7	6	3,298	.31	.075	.41	.087	.30	.50	.083	.39
8	51	44,637	.25	.064	.35	.078	.25	.51	.075	.42
9	34	7,736	.30	.092	.40	.091	.28	.51	.097	.38
10	5	1,340	.31	.067	.42	.052	.35	.48	.041	.43
11	5	1,242	.37	.082	.46	.076	.37	.60	.045	.54
12	3	301	.29	.126	.40	.123	.24	.50	.018	.47
13	39	10,313	.40	.092	.45	.085	.34	.53	.067	.45
14	17	38,909	.49	.034	.57	.035	.52	.64	.053	.58
15	23	7,786	.32	.099	.44	.114	.29	.46	.091	.34
16	29	13,730	.34	.080	.41	.082	.30	.48	.087	.37
17	10	2,796	.39	.080	.48	.077	.39	.56	.000	.56
20	13	8,564	.31	.084	.43	.107	.29	.48	.096	.36
Total	379	255,490								
Weighted means[c]			.32	.076	.41	.080	.31	.51	.072	.42

[a]Please refer to Table 1 for full names of composites.

[b]Credibility value.

[c]Means weighted according to sample size.

Table 6

Statistical Summary for BTB Composites for Rating Families Based on the Navy Basic Test Battery (BTBFAM)

Composite and Occupational Group	N Samples	N Students	Observed Validity Mean	Observed Validity SD	Validity Generalization Based on MCV Mean	Validity Generalization Based on MCV SD	Validity Generalization Based on MCV 90% CV[b]	Validity Generalization Based on ICV Mean	Validity Generalization Based on ICV SD	Validity Generalization Based on ICV 90% CV[b]
GT										
1	101	80,682	.48	.109	.64	.140	.46	.62	.114	.47
2	146	104,417	.43	.090	.59	.116	.44	.60	.088	.49
3	84	62,824	.36	.076	.60	.115	.45	.62	.072	.52
4	36	13,348	.45	.133	.58	.163	.37	.56	.138	.38
Total	367	261,271								
Weighted means[c]			.43	.096	.60	.127	.44	.61	.096	.48
MECH										
1	71	68,652	.33	.111	.47	.150	.27	.49	.131	.32
2	89	75,425	.43	.079	.63	.109	.49	.61	.085	.50
3	53	46,840	.39	.072	.58	.097	.45	.59	.061	.52
4	26	11,928	.29	.106	.41	.140	.23	.44	.125	.28
Total	239	202,845								
Weighted means[c]			.38	.090	.55	.122	.39	.55	.098	.43
ELEC										
1	56	59,051	.46	.110	.63	.147	.44	.57	.126	.41
2	69	56,770	.48	.093	.67	.124	.51	.61	.110	.47
3	53	46,840	.50	.095	.81	.150	.62	.71	.095	.59
4	15	7,132	.45	.103	.62	.132	.45	.54	.093	.42
Total	193	169,793								
Weighted means[c]			.48	.099	.69	.138	.52	.62	.109	.48
CLER										
1	105	81,614	.40	.101	.49	.117	.34	.56	.112	.41
2	145	103,254	.33	.074	.42	.084	.31	.52	.076	.42
3	82	62,307	.26	.069	.36	.084	.25	.52	.068	.43
4	33	11,789	.33	.101	.45	.120	.29	.47	.118	.32
Total	365	258,964								
Weighted means[c]			.33	.083	.43	.097	.30	.53	.088	.41

[a]Please refer to Table 1 for full names of composites.

[b]Credibility value.

[c]Means weighted according to sample size.

Table 7

Average Within-Group Estimated Standard Deviations (SDs) of True Validities for BTB Composites by Job Grouping Strategy and by Random Grouping

Rating Family	SDs for BTB Composites									
	GT		MECH		ELEC		CLER		Means	
NOHFAM[a]	.091	(.092)	.091	(.110)	.112	(.116)	.078	(.080)	.093	(.100)
CLASSFAM[b]	.082	(.085)	.079	(.093)	.090	(.094)	.072	(.073)	.081	(.086)
BTBFAM[c]	.096	(.097)	.097	(.118)	.109	(.119)	.088	(.088)	.098	(.105)
All ratings	.098		.118		.125		.093		.109	

Note. Random groupings are indicated by parentheses.

[a]Navy Occupational Handbook (U.S. Navy, 1966).

[b]Official classification manual (U.S. Navy, 1975).

[c]BTB composites used for "A" school assignment.

Table 8

Average Within-Group Estimated Variance of True Validities for Substantive and Random Grouping Systems as Percent of Variance for all Navy Ratings

Rating Family	Percent of Variance by BTB Composite														
	GT			MECH			ELEC			CLER			Means		
	1	2	Difference	1	2	Difference	1	2	Difference	1	2	Difference	1	2	Difference
NOHFAM[a]	86	88	2	59	87	28	80	86	6	70	74	4	74	84	10
CLASSFAM[b]	70	75	5	45	62	17	52	56	4	60	62	2	57	64	7
BTBFAM[c]	96	98	2	67	100	33	76	91	15	90	90	0	82	95	13

[a]Navy Occupational Handbook (U.S. Navy, 1966).

[b]Official classification manual (U.S. Navy, 1975).

[c]BTB composites used for "A" school assignment.

percent reduction in variance observed for the NOHFAM, 16 percent was due to chance. Thus, 61 percent (16/26 X 100) of the 26 percent reduction in variance observed for the NOHFAM rating family can be attributed to chance. The 10 in the Difference column refers to the amount of reduction in variance that is attributable to the NOHFAM rating family. Thus, 38 percent (10/26 X 100) of the 26 percent reduction in variance observed for the NOHFAM is attributable to the NOHFAM rating system.

A reduction of variance of 43 percent (100% - 57%) and 18 percent (100% - 82%) was observed for the CLASSFAM and BTBFAM for values averaged across the BTB composites. This is the reduction in variance attributable to the CLASSFAM and BTBFAM rating systems as compared to the ungrouped Navy ratings. Of these reductions in variance, 36 percent and 5 percent respectively were due to chance. Thus 84 percent (36/43 x 100) of the reduction in variance for CLASSFAM and 28 percent (5/18 x 100) of the reduction in variance for BTB can be attributed to chance. The Difference columns for CLASSFAM and BTBFAM show 7 and 13 respectively, the amount of reduction in variance attributable to these two rating family systems. The related percent reductions in variance attributable to CLASSFAM and BTBFAM are 16 percent (7/43 x 100) and 72 percent (13/18 x 100) respectively.

As can be seen in Table 7, the CLASSFAM rating system reduced variability of validity for each of the four BTB composites more than either the NOHFAM or BTBFAM rating systems. The CLASSFAM system, averaged across composites in Table 8, produced the greatest reduction in variance (43%) of all Navy ratings. However, 84 percent of this was attributable to chance. Overall, the three rating family systems showed no substantial differences in reducing variability of validity.

CONCLUSIONS AND RECOMMENDATIONS

The results of this study support the broad generalizability of validity data and the limited effect that the three systematically formed rating families play in moderating the variability of validity. These findings were similar to Pearlman and Schmidt's (1981; Pearlman, 1982), who investigated 15 systematically formed rating families. Pearlman and Schmidt (1981) concluded that, ". . . simple, rational groupings based on the general content structure of jobs are equally useful as grouping derived by more complex, time consuming, and expensive methods." (p. 11)

It is recommended that a data base of validity coefficients for ASVAB selector composites be developed that is similar to the BTB data base. This information will be used to (1) provide estimates of the validities of the ASVAB selector composites at the time a new ASVAB form is introduced and before it can be validated, (2) evaluate the need to revalidate a selector composite for a Navy rating when there is a curriculum change, a change in method of instruction, or the creation of a new rating, and (3) suggest ways of combining criterion data from small Class "A" schools with larger schools to provide samples of sufficient size for validation.

REFERENCES

Booth-Kewley, S. (February 1985). An empirical comparison of the accuracy of univariate and multivariate corrections for range restriction. (NPRDC Tech. Rep. 85-19). San Diego: Navy Personnel Research and Development Center. (AD-A153 071)

Guilford, J. P. (1965). Fundamental statistics in psychology and education (4th ed.). New York: McGraw-Hill.

Naval Personnel Research Activity. (July 1961). Basic Test Battery validity reports for 89 class "A" school samples (1957-1958). San Diego: Author.

Outerbridge, A. N. (August 1979). A survey of test validation study costs (Tech. Memo. 79-18). Washington, DC: Office of Personnel Management, Personnel Research and Development Center.

Pearlman, K. (1982). The Bayesian approach to validity generalization: A systematic examination of the robustness of procedures and conclusions. Unpublished doctoral dissertation, George Washington University, Washington, DC.

Pearlman, K., & Schmidt, F. L. (August 1981). Effects of alternate job grouping methods on selection procedure validity. In E. L. Levine (Chair), Job Analysis/Job Formulas: Current Perspectives on Research and Application. Symposium conducted at the annual meeting of the American Psychological Association, Los Angeles.

Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. Journal of Applied Psychology, 65, 373-406.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. Journal of Applied Psychology, 62, 529-540.

Schmidt, F. L., Hunter, J. E., & Caplan, J. K. (1981). Validity generalization results for two job groups in the petroleum industry. Journal of Applied Psychology, 66, 261-273.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1981). Task differences as moderators of aptitude test validity in selection: A red herring. Journal of Applied Psychology, 66, 166-185.

Schmidt, F.L., Hunter, J. E., & Pearlman, K. (1982). Progress in validity generalization: Comments on Callender and Osborn and further developments. Journal of Applied Psychology, 67, 835-845.

Schmidt, F. L., Hunter, J. E., Pearlman, K., & Shane, G. S. (1979). Further tests of the Schmidt-Hunter Bayesian validity generalization procedure. Personnel Psychology, 32, 257-281.

Swanson, L. (1977). Validity data on form 7 of USN Basic Test Battery. Unpublished data. San Diego: Navy Personnel Research and Development Center.

Swanson, L., & Foley, P. (November 1982). Development of Armed Forces Qualification Test (AFQT) deliberate failure keys for Armed Services Vocational Aptitude Battery (ASVAB) forms 8, 9, and 10. (NPRDC Spec. Rep. 83-4). San Diego: Navy Personnel Research and Development Center.

Thomas, P. J. (April 1970). Validation of Form 7 of the USN Basic Test Battery in 108 class "A" schools (Tech. Bul. STB 70-7). San Diego: Naval Personnel and Training Research Laboratory.

Thomas, P. J., & Thomas, E. D. (August 1965). School validation of the Basic Test Battery and of other classification tests (1962-1963) (Tech. Bul. STB 66-8). San Diego: Naval Personnel Research Activity.

Thomas, P. J., & Thomas, E. D. (October 1967). Validation of form 7 of the USN Basic Test Battery in 91 class "A" schools (Tech. Bul. STB 68-4). San Diego: Naval Personnel Research Activity.

Thorndike, R. L. (1982). Applied psychometrics. Boston: Houghton Mifflin.

U.S. Department of the Navy. (September 1975). Manual of Navy enlisted manpower and personnel classification and occupational standards. Washington, DC: Author.

U.S. Department of the Navy. (1966). U.S. Navy occupational handbook. Washington, DC: Author.

APPENDIX A

TESTS IN BASIC TEST BATTERY (BTB), FORMS 6 AND 7

Appendix A

Tests in Basic Test Battery (BTB), Forms 6 and 7

Test	Content	Form 6	Form 7
General classification test (GCT)	40 sentence completion	X	X
	60 verbal analogies	X	X
Arithmetic (ARI)	30 arithmetic reasoning	X	X
	20 arithmetic computation	X	
Mechanical (M)	50 tool knowledge	X	X
	50 mechanical comprehension	X	X
Clerical (CLER)	210 number matching (5 to 9 digits)	X	X
Shop practices (SP)	30 shop information and knowledge		X
Electronics technician	20 mathematics (including algebra)	X	X
Selection test (ETST)	20 general science (mostly physics)	X	X
	10 shop practice	X	
	15 electricity knowledge	X	X
	15 radio knowledge	X	X

APPENDIX B

NUMBER OF BTB VALIDATION STUDIES BY RATING

Appendix B

Number of BTB Validation Studies by Rating

Rating	Course Data Processing Code	Number of Studies
Steward (SD)	6000	3
Quartermaster (QM)	6002	8
Signalman (SM)	6007	8
Sonar Technician (STG) (Surface)	6015	2
Sonar Technician (STS) (Submarine)	6017	3
Gunner's Mate Technician (GMT)	6025	4
Fire Control Technician (FT)	6028	13
Torpedoman's Mate (TM)	6034	6
Torpedoman's Mate (TM)	6036	1
Missile Technician (MT)	6040	2
Mineman (MN)	6041	3
Instrumentman (IM)	6046	4
Opticalman (OM)	6047	4
Cryptologic Technician (CTA) (Administrative)	6051	5
Cryptologic Technician (CTM) (Maintenance)	6052	2
Cryptologic Technician (CTO) (Communications)	6053	5
Cryptologic Technician (CTR) (Collection)	6054	2
Yeoman (YN)	6058	13
Storekeeper (SK)	6060	10
Disbursing Clerk (DK)	6062	10
Journalist (JO)	6063	2
Draftsman (DM)	6064	2
Machinist's Mate (MM)	6066	4
Engineman (EN)	6067	5
Machinery Repairman (MR)	6068	5
Boiler Technician (BT)	6069	5
Electrician's Mate (EM)	6071	12
Interior Communications Electrician (IC)	6073	11
Hospitalman (HM)	6085	12
Dental Technician (DT)	6086	5
Postal Clerk (PC)	6090	4
Gunner's Mate (GMM) (Missles)	6096	3
Cryptologic Technician (CTT) (Technical)	6099	1
Cryptologic Technician (CTI) (Interpretive)	6100	2
Personnelman (PN)	6102	7
Communications Yeoman (CYN)	6105	3
Shipfitter (SF)	6106	8
Fire Control Technician (FT)	6109	4
Gunner's Mate-Guns (GMG)	6110	2
Gunner's Mate (GM)	6115	2
Gunner's Mate-Guns (GMG)	6117	2
Fire Control Technician (FT)	6118	1
Damage Controlman (DC)	6120	5
Commissaryman (CS)	6125	4
Electronics Technician (ET)	6130	4
Electronics Technician (ET)	6134	6

Appendix B (Continued)

Rating	Course Data Processing Code	Number of Studies
Electronics Technician-Radar (ETR)	6136	11
Electronics Technician-Communications (ETN)	6138	12
Data Systems Technician (DS)	6139	1
Operations Specialist (OS)	6142	12
Radioman (RM)	6144	2
Polaris Electronic (PE)	6146	2
Engineering Aid (EA)	6147	1
Construction Electrician (CE)	6148	7
Construction Mechanic (CM)	6149	8
Builder (BU)	6150	7
Steelworker (SW)	6151	5
Utilitiesman (UT)	6152	6
Equipment Operator (EO)	6159	6
Sonar Technician (ST)	6160	1
Torpedoman's Mate (TM)	6169	11
Gunner's Mate Technician (GMT)	6170	1
Data Processing Technician (DP)	6171	2
Aviation Machinist's Mate (ADR) (Reciprocating)	6502	12
Aviation Electronics Technician (ATN) (Navigation)	6503	6
Aviation Electronics Technician (ATR) (Radar)	6504	4
Aviation Antisubmarine Warfare Operator (AX)	6505	5
Aviation Ordnanceman (AO)	6506	5
Aviation Fire Control Technician (AQB) (Bomb Director)	6507	2
Aviation Fire Control Technicial (AQF)	6508	5
Air Traffic Controller (AC)	6509	6
Aviation Boatswain's Mate, Fuels (ABF)	6512	3
Aviation Boatswain's Mate, Launch (ABE) (Launch and Recovery)	6513	3
Aviation Electrician's Mate (AE)	6515	6
Aviation Structural Mechanic (AME) (Safety Equipment)	6516	4
Aviation Structural Mecahnic (AMH) (Hydraulics)	6517	5
Aviation Structural Mechanic (AMS) (Structures)	6518	6
Aircrew Survival Equipmentman (PR)	6519	6
Aerographer's Mate (AG)	6520	5
Tradevman (TD)	6521	5
Aviation Storekeeper (AK)	6522	5
Photographer's Mate (PH)	6523	6
Aviation Electronics Technician (ATW) (Airborne CIC Operator)	6526	1
Aviation Boatswain's Mate (ABH) (Aircraft Handling)	6527	5
Aviation Maintenance Administrationman (AZ)	6528	3
Photographic Intelligence Man (PT)	6529	2
Aviation Support Equipment Technician (AS)	6530	2
Avionics Fundamentals (AF)	6533	8
Aviation Fire Control Technician (AQ)	6535	3
Anti-Submarine Warfare Operator (AW)	6537	2

APPENDIX C

EXPLANATION AND WORKED EXAMPLES FOR TWO ESTIMATION PROCEDURES

Appendix C

Explanation and Worked Examples for Two Estimation Procedures

Both the mean corrected validities (MCV) and individually corrected validities (ICV) procedures provide estimates of the mean and SD of distributions of true validity coefficients. In the MCV procedure, the mean of each distribution was corrected for range restriction and criterion unreliability. In the ICV procedure, the validities were individually corrected for these sources of error. An explanation of the validity generalization procedure used and a worked example are provided within each of the following sections. The data used for the worked examples (see Table C-1) is for the QM rating, which is the first rating shown in Table 2 in the text.

Table C-1

Statistical Data for Worked Example of the MCV Procedure for the QM Rating

Validation Study	Subjects (N)	Observed $\underline{r}$	SD	n(r)	$n(r^2)$	$\frac{n(1-r^2)^2}{n-1}$	n(SD)
1	558	.34	10.75	189.72	64.50	.7836	5998.50
2	83	.35	12.80	29.05	10.16	.7794	1062.40
3	346	.34	10.12	117.64	39.99	.7845	3501.52
4	496	.53	14.15	262.88	139.32	.5181	7018.40
5	1156	.41	10.93	473.96	194.32	.6926	12635.08
6	144	.55	13.44	79.20	43.56	.4899	1935.36
Total	2783			1152.45	491.85	4.0481	32151.26

Mean Corrected Validities (MCV)

The first procedure used distributions of validity coefficients that had not been corrected for restriction in range. The steps involved were as follows:

1. The sample-sized weighted mean ($\bar{r}$) and variance (σ^2 total) of each distribution of observed validities were computed.

$$\bar{r} = 1152.45/2783 = .414$$

$$(\sigma^2 \text{ total}) = (491.85/2{,}783) - [(1152.45/2{,}783)]^2 = .00526$$

2. The mean of each distribution was corrected for range restriction using Thorndike's (1982) case A formula. V equals the ratio of the unrestricted to the restricted standard deviation. V for the current problem was (13.75/11.55 = 1.19). The restricted SD

of 11.55 (32151.26/2783) is the weighted average of restricted SDs for the six validity studies. The unrestricted SD of 13.75 for the CLER composite was developed from variances of subtests and their intercorrelation provided in Thomas and Thomas (1965) for BTB Form 6 and Thomas (1967) for Form 7. The technique (variance and sum of a composite) used for calculating the composite SD is provided in Guilford (1965, p. 541). The unrestricted SDs for the MECH, ELEC, and CLER composites are 23.30, 27.90, and 13.75.

$$R(\text{corrected for range restriction}) = \bar{r}V/\sqrt{1 - \bar{r}^2 + \bar{r}^2 V^2}$$

$$R = .41(1.19)/\sqrt{1 - (.41)^2 + (.41)^2 (1.19)^2} = .4761$$

An estimate of mean true validity was obtained by correcting R for attenuation due to criterion unreliability, using an assumed criterion reliability of .90.

$$\text{Mean true validity} = .4761/\sqrt{.90} = .5018$$

3. The sampling error variance (σ^2_N) of each distribution of observed validities was computed using the following formula:

$$\sigma^2_N = \frac{\sum \left[\frac{n(1 - r^2)^2}{n - 1}\right]}{N}$$

where r = an observed validity coefficient and n = the sample size associated with r.

$$\sigma^2_N = 4.0481/2{,}783 = .00145$$

4. The residual standard deviation is equal to

$$\sqrt{\sigma^2 \text{ total} - \sigma^2_N}$$

$$\sqrt{.00526 - .00145} = .06172$$

5. The residual standard deviation was multiplied by the ratio (estimated mean "true" validity/mean observed validity) to provide an estimate of the standard deviation of the distribution of true validities (corrected for sampling error).

$$\text{SD true validities} = (.5018/.414)\,(.06172) = .0748$$

6. 90-percent-credibility values were computed for each distribution by multiplying the estimate of the SD of the distribution of "true" validities by 1.2816, that point on the abcissa of the normal curve below which 90 percent of the area lies, and subtracting this value from the estimated mean true validity.

$$(.0748)\,(1.2816) - .5018 = .41$$

Individually Corrected Validities (ICV)

This approach used distributions of validity coefficients that had previously been individually corrected for restriction in range. Qualifying scores for Class "A" school assignments vary across schools, but all have the effect of reducing the observed composite validities from what they would be for a full range population. For this reason, the observed validities are commonly corrected for restriction in range and the corrected values are typically reported in validity studies. The steps involved are as follows:

1. Each range-restriction corrected validity was corrected for attenuation due to criterion unreliability, producing an estimate of "true" validity shown in Table C-2. The assumed criterion reliability was .90.

Table C-2

Statistical Data for Worked Example of the ICV Procedure for the QM Rating

Validation Study	Subjects N	True $\underline{r}$	n(r)	n(r^2)
1	558	.56	312.48	174.99
2	83	.51	42.33	21.59
3	346	.61	211.06	128.75
4	496	.62	307.52	190.66
5	1156	.50	578.00	289.00
6	144	.61	87.84	53.58
Total	2783		1539.23	858.57

2. The sample-size weighted mean ($\bar{r}$) and variance (σ^2 total) of each distribution of estimated "true" validities were computed.

$$\bar{r} = 1539.23/2783 = .5531$$

$$(\sigma^2 \text{ total}) = (858.57/2783) - [(1539.23/2783)]^2 = .00260$$

3. The sampling error variance of each true validity distribution was estimated by multiplying the average sample size weighted sampling error of observed validities (σ^2_N= .00145) by the squared ratio (estimated mean "true" validity/mean observed validity). See (Pearlman et al., 1980, p. 404) for further explanation.

$$\sigma^2_N = .00145\ (.5531/.414) = .00259$$

4. The variance of the estimated true validities was corrected for sampling error variance (σ^2 total $-\sigma^2_N$). The square root of this corrected variance is the SD of the Bayesian prior distribution.

$$\text{SD(prior)} = \sqrt{.00260 - .00259} = .00316$$

5. The 90-percent-credibility value for each distribution was calculated by multiplying the SD (prior) by 1.2816 and subtracting this value from the sample size weighted mean of estimated true validity:

$$90\%\ \text{CV} = .00316(1.2816) - .5531 = .55.$$

APPENDIX D

NAVY "A" SCHOOLS AND RATINGS USED FOR VALIDITY GENERALIZATION OUTCOMES

Appendix D

Navy "A" Schools and Ratings used for Validity Generalization Outcomes

School and Ratings	Composite
Quartermaster (QM)	CLER
Signalman (SM)	CLER
Fire control technician (FT)	ELEC
Torpedoman's mate (TM)	GT
Yeoman (YM)	CLER
Storekeeper (SK)	CLER
Disbursing clerk (DK)	GT
Electrician's mate (EM)	MECH
Interior communications electrician (IC)	MECH
Hospitalman (HM)	GT
Electronics technician (ET)	ELEC
Electronics technician (ETR)	ELEC
Electronics technician (ETN)	ELEC
Operations specialist (OS)	GT
Torpedoman's mate (TM)	GT
Photographer's mate (PH)	GT

APPENDIX E

ALTERNATIVE OCCUPATIONAL GROUPINGS FOR NAVY RATINGS

Appendix E

Alternative Occupational Groupings for Navy Ratings

	Occupational Categories	Rating Codes
	Groupings Based on the *Navy Occupational Handbook* (NOH)	
1.	Deck group	QM, SM, ST
2	Ordnance group	FT, GM, MN, MT, TM
3.	Electronics and precision instruments group	DS, ET, IM
4.	Administrative and clerical group	CT, DK, DP, JO, MS(CS), PC, PN, RM, SK, YN
5.	Engineering and hull group	BT, EM, EN, IC, MM, MR
6.	Construction group	BU, CE, CM, EA, EO, SW, UT
7.	Aviation group	AB, AC, AD, AE, AG, AK, AM, AO, AQ, AS, AT, AW, AX, AZ, PH, PR, TD
8 & 9.	Miscellaneous, medical, dental and steward group	DM, DT, HM, MS(SD)
	Groupings Based on the Classification Manual (CLASSMAN)	
1.	General seamanship	SM
2.	Ship operations	QM
3.	Marine engineering	BT, EM, EN, IC, MM
4.	Ship maintenance	IM, MR, OM
5.	Aviation maintenance/weapons	AD, AE, AM, AO, AQ, AT, AX, AZ, PR
6.	Aviation ground support	AB, AS
7.	Air traffic control	AC
8.	Weapons control	ET, FT
9.	Ordnance systems	GM, MN, MT, TM
10.	Sensor operations	ST
11	Weapons system support	TD
12.	Data systems	DP, DS
13.	Construction	BU, CE, CM, EA, EO, SW, UT
14.	Health care	DT, HM
15.	Administration	PC, PN, YN
16.	Logistics	AK, DK, MS, SK
17.	Media	DM, JO, PH
20.	Cryptology	CT
23.	Meteorology	AG
24.	Aviation sensor operations	AW
	Groupings Based on the Navy Basic Test Battery (BTBFAM)	
1.	GCT + ARI	AC, AG, AK, AZ, CT (except CTA, CTI, CTM) DK, DM, DP, DT, EA, HM, MS, PH, PN, RM, SK, ST, TM
2.	GCT + MECH + SP	AB, AD, AE, AM, AO, AS, BT, BU, CE, CM, EM, EN, EO, GM, IC, IM, MM, MN, MR, OM, PR, SW, UT
3.	ARI + 2 ETST	AQ, AT, AW, AX, CTM, DS, ET, FT, MT, TD
4.	GCT + CLER	CTA, CTI, JO, QM, SM, YN

DISTRIBUTION LIST

Deputy Assistant Secretary of the Navy (Manpower)
Chief of Naval Operations (OP-01B6), (OP-01B7) (2), (OP-11), (OP-135C), (OP-14), (OP-140F2), (OP-987H)
Office of Naval Technology (ONT-0722)
Commander, Naval Military Personnel Command (NMPC-00), (NMPC-013D), (NMPC-4)
Commander, Navy Recruiting Command (Code 20)
Commanding Officer, Naval Aerospace Medical Institute (Code 012), Pensacola
Chief of Naval Research (Code 270), (Code 440), (Code 442), (Code 442PT)
Chief of Naval Education and Training (Code 00A), (Code N-21)
Chief of Naval Technical Training (Code N-6)
Commanding Officer, Naval Technical Training Center, Corry Station (Code 101B), Pensacola
Commanding Officer, Naval Training Equipment Center (Technical Library) (5), (Code 10)
Officer In Charge, Navy Occupational Development and Analysis Center, Washington, DC
Commandant of the Marine Corps (MPI-20)
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria (PERI-POT-I), (PERI-SZ)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base, TX (AFHRL/MO Manpower and Personnel Division (2), (Scientific and Technical Information Officer, TSRL/Technical Library FL 2870)
Commander, OPSTNGDIV Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/OT)
Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (AFHRL/LR-TDC)
Commanding Officer, U.S. Coast Guard Research and Development Center, Avery Point
Superintendent, Naval Postgraduate School
Director of Research, U.S. Naval Academy
Canadian Forces Personnel, Applied Research Unit, Canada
DPSRSC, National Defense Headquarters, Canada
Ministry of Defense, Senior Psychologist, England (Naval)
D. Dennison, Army Personnel Research Establishment, Personnel Psychological Division, England (2)
Science 3 (RAF), Lacon House, England
1 Psychological Research Unit, NBH 3-44, Australia
Directorate of Psychology - AF, Department of Defense, (Air for CE) Australia
Navy Psychology, Australia (2)
Defense Psychology Unit, Defense HQ (2)
Dr. Regan
Defense Technical Information Center (DDAC) (2)

DEPARTMENT OF THE NAVY
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
(CODE___)
SAN DIEGO, CA 92152-6800

0142